SEQUENTIAL METHODS IN STATISTICS
BANACH CENTER PUBLICATIONS, VOLUME 16
PWN–POLISH SCIENTIFIC PUBLISHERS
WARSAW 1985

AN OPTIMALITY-EQUATION FOR DISCRETE STOCHASTIC DECISION PROBLEMS WITH GENERAL SETS OF ADMISSIBLE STRATEGIES

P. HEINZ MÜLLER

Section of Mathematics, Technical University of Dresden, Dresden, G.D.R.

0. Introduction

Considered are stochastic decisions with discrete time space and countable state space. For rather general classes of admissible strategies the corresponding optimality equation is formulated. Included are especially problems of optimal stopping.

1. Description of the dynamic system

(a) *Space of time*

$$T_N = \begin{cases} \{0, 1, \ldots, N\} & \text{if } N < \infty, \\ \{0, 1, \ldots\ \} & \text{if } N = \infty. \end{cases}$$

We call N the *horizon* and the elements of T_N *points of time* or *steps*.

(b) *Space of states* $\mathcal{S}$, $\mathcal{S}$ is supposed to be countable and $\mathcal{S} \neq \emptyset$, elements of $\mathcal{S}$: s, s_n.

(c) *Space of decisions* $\mathcal{D}$, $\mathcal{D} \neq \emptyset$, elements of $\mathcal{D}$: d, d_n.

$h_n := (s_0, d_0, s_1, d_1, \ldots, d_{n-1}, s_n)$, $n \geqslant 0$ is called a *n-history*,

$$\mathcal{H}_n := \underbrace{(\mathcal{S} \times \mathcal{D} \times \mathcal{S} \times \mathcal{D} \times \ldots \times \mathcal{D} \times \mathcal{S})}_{(2n+1)\text{ factors}} \text{ set of } n\text{-histories.}$$

(d) *Initial distribution* $p_0(\cdot)$, probability distribution on $\mathcal{S}$.

(e) *Law of motion* $\mathfrak{p} = (p_n)$, $n = 1, 2, \ldots, p_n$: probability distribution on $\mathcal{S}$ of the form $p_n(\cdot | h, d)$, $h \in \mathcal{H}_n$, $d \in \mathcal{D}$.

(f) *Cost structure* $\mathfrak{k} = (k_n)$, $n = 1, 2, \ldots, k_n$, cost function of the form $k_n | \mathcal{H}_n \to \boldsymbol{R}$.

(g) *Decision function* for time n: $\delta_n|\mathscr{H}_n \to \mathscr{D}$, $0 \leqslant n < N$.

(h) *Strategy* $\vartheta = (\delta_n)$, $0 \leqslant n < N$.

(i) *Set of admissible strategies* Θ.

By the entity DS: $(N, \mathscr{S}, \mathscr{D}, p_0, \mathfrak{p}, \mathfrak{k}, \Theta)$ a *dynamical system* is defined.

2. Notation

Each vector of the form $\mathfrak{s}_n := (s_0, s_1, \ldots, s_n)$, $\mathfrak{s}_n \in \mathscr{S}^{n+1}$, is called a *state-n-history*. In case $N = \infty$ we write $\mathfrak{s}_N := (s_0, s_1, \ldots)$. By the states $s_0, \ldots, s_n$ contained in a n-history $h_n = (s_0, d_0, \ldots, d_{n-1}, s_n)$ we may form the vector of state-n-history $\mathfrak{s}_n(h_n) := (s_0, \ldots, s_n)$. To a n-history $h_n = (s_0, d_0, \ldots, d_{n-1}, s_n)$ or a state-n-history $\mathfrak{s}_n = (s_0, \ldots, s_n)$ we define for $0 \leqslant k \leqslant n$ the *k-sections* $a_k h_n := (s_0, d_0, \ldots, d_{k-1}, s_k)$ and $a_k \mathfrak{s}_n := (s_0, \ldots, s_k)$. Immediately we understand the abbreviated notation $h_{n+1} = (h_n, d_n, s_{n+1})$, $n \geqslant 0$. A n-history $h_n = (s_0, d_0, \ldots, d_{n-1}, s_n)$ and a strategy $\vartheta = (\delta_k)$ are called *compatible* – in symbols $h_n \sim \vartheta$ or $\vartheta \sim h_n$ – if $\delta_k(a_k h_n) = d_k$ for all $k = 0, 1, \ldots, n-1$. To given $\mathfrak{s}_n = (s_0, s_1, \ldots, s_{n-1}, s_n)$ and $\vartheta = (\delta_k)$ we can construct by $d_0 := \delta_0(s_0)$, $d_1 := \delta_1(s_0, \delta_0(s_0)), \ldots, d_{n-1} := (s_0, \delta_0(s_0), \ldots, \delta_{n-1}(s_0, \delta_0(s_0), \ldots, s_{n-1}))$ the n-history $(s_0, d_0, \ldots, d_{n-1}, s_n) =: h_n^\vartheta(\mathfrak{s}_n)$, which is characterized by $h_n \sim \vartheta$, $\mathfrak{s}_n = \mathfrak{s}_n(h_n)$, $\delta_k(a_k h_n) = d_k$ for $0 \leqslant k \leqslant n-1$. Finally for $\vartheta = (\delta_k)$ we put

$$p_{n,\vartheta}(\cdot|\mathfrak{s}_{n-1}) := p_n(\cdot|h_{n-1}^\vartheta(\mathfrak{s}_{n-1}), \delta_{n-1}(h_{n-1}^\vartheta(\mathfrak{s}_{n-1}))) \tag{1}$$

and

$$k_{n,\vartheta}(\mathfrak{s}_n) := k_n(h_n^\vartheta(\mathfrak{s}_n)). \tag{1'}$$

3. The controlled stochastic process

In the usual way we define by the theorem of Ionescu-Tulcea to each $\vartheta \in \Theta$ a stochastic process (X_n), $n = 0, 1, \ldots$ on a probability space $[\Omega, \mathscr{F}, \boldsymbol{P}_\vartheta]$ with $\Omega = \bigtimes_0^N \mathscr{S}$, $\mathscr{F} = \bigotimes_0^N \mathfrak{P}(\mathscr{S})$, $X_n(\omega) = s_n$ for $\omega = (s_0, s_1, \ldots) \in \Omega$ and the probability measure $\boldsymbol{P}_\vartheta$ defined by

$$\begin{aligned} &\boldsymbol{P}_\vartheta(X_0 = s_0) := p_0(s_0), \qquad s_0 \in \mathscr{S}, \\ &\boldsymbol{P}_\vartheta(X_n = s_n|X_0 = s_0, \ldots, X_{n-1} = s_{n-1}) := p_{n,\vartheta}(s_n|\mathfrak{s}_{n-1}), \end{aligned} \tag{2}$$

$$(s_0, s_1, \ldots, s_{n-1}) = \mathfrak{s}_{n-1} \in \mathscr{S}^n, \qquad n = 1, 2, \ldots$$

By means of the notation $\mathfrak{S}_n := (X_0, X_1, \ldots, X_n)$ for the random state-n-history we have in equivalence to (2) the relation

$$\boldsymbol{P}_\vartheta(\mathfrak{S}_n = \mathfrak{s}_n) = p_0(s_0) \prod_{\nu=1}^n p_{\nu,\vartheta}(s_\nu|\mathfrak{s}_{\nu-1}),$$

$$\mathfrak{s}_n = (s_0, s_1, \ldots, s_n), \qquad \mathfrak{s}_\nu = a_\nu \mathfrak{s}_n, \qquad 1 \leqslant \nu \leqslant n. \tag{2'}$$

We design by $X_\vartheta := \{(X_n), n = 0, 1, \ldots | [\Omega, \mathscr{F}, \boldsymbol{P}_\vartheta]\}$ the *stochastic process controlled by* ϑ, $\vartheta \in \Theta$.

A state-n-history $\mathfrak{s}_n$, $n \geqslant 0$, is called *essential* for ϑ, if $\boldsymbol{P}_\vartheta(\mathfrak{S}_n = \mathfrak{s}_n) > 0$.

4. Formulation of the problem

In dependance of the random state-n-history $\mathfrak{S}_n$ we define (compare (1')) the *random cumulative cost*

$$k_{n,\vartheta}(\mathfrak{S}_n) =: K_{n,\vartheta}, \qquad n = 1, 2, \ldots \tag{3}$$

and thereby the *aim function*

$$v_\vartheta = \begin{cases} \boldsymbol{E}_\vartheta(K_{n,\vartheta}) & \text{if } N < \infty, \\ \overline{\lim\limits_{n\to\infty}}\, \boldsymbol{E}_\vartheta(K_{n,\vartheta}) & \text{if } N = \infty, \ \vartheta \in \Theta. \end{cases} \tag{4}$$

(The existence of the corresponding expectation values is supposed.) A strategy $\vartheta^* \in \Theta$ is called *optimal*, if

$$v_{\vartheta^*} = \inf_{\vartheta \in \Theta} v_\vartheta := v^*. \tag{5}$$

5. Notation

We define by

$$c_n(h_n) := k_n(h_n) - k_{n-1}(a_{n-1}\, h_n), \qquad n = 1, 2, \ldots, k_0 :\equiv 0 \tag{6}$$

step cost $c_n | \mathscr{H}_n \to \boldsymbol{R}$ for time n and in dependance of $h_N = (s_0, d_0, \ldots, s_n, d_n, \ldots)$ by

$$r_n(h_N) := k_N(h_N) - k_{n-1}(a_{n-1}\, h_N), \qquad n = 1, 2, \ldots, \tag{7}$$

rest cost $r_n | \mathscr{H}_n \to \bar{\boldsymbol{R}}$ for time n.

Given a strategy ϑ and a state-history $\mathfrak{s}_n$ or $\mathfrak{s}_N$ we use in connection with $h_n = h_n^\vartheta(\mathfrak{s}_n)$ and $h_N = h_N^\vartheta(\mathfrak{s}_N)$ the notation

$$c_{n,\vartheta}(\mathfrak{s}_n) := c_n(h_n), \qquad r_{n,\vartheta}(\mathfrak{s}_N) := r_n(h_N), \tag{8}$$

and in connection with this we obtain by

$$C_{n,\vartheta} | \mathfrak{S}_n \to c_{n,\vartheta}(\mathfrak{S}_n), \qquad R_{n,\vartheta} | \mathfrak{S}_N \to r_{n,\vartheta}(\mathfrak{S}_N) \tag{9}$$

the random step and rest cost at time n.

The *mean step cost* for time $n = 0, 1, 2,,,$ are because of (1), (6), (8) defined by

$$E(C_{n+1,\vartheta} | \mathfrak{S}_n = \mathfrak{s}_n) = \sum_s p_{n+1,\vartheta}(s|\mathfrak{s}_n)\, c_{n+1,\vartheta}(\mathfrak{s}_n, s) =: \mathfrak{C}_{n,\vartheta}(\mathfrak{s}_n), \tag{10}$$

and for $\vartheta = (\delta_k) \sim h_n$, $\mathfrak{s}_n = \mathfrak{s}_n(h_n)$, $d_n = \delta_n(h_n)$, $n = 0, 1, 2, \ldots$ we have (compare (1), (8))

$$\mathfrak{C}_{n,\vartheta}(\mathfrak{s}_n) = \sum_s p_{n+1}(s|h_n, d_n) c_{n+1}(h_n, d_n, s) =: \mathfrak{C}_n(h_n, d). \tag{10'}$$

We define (*conditional*) *mean rest cost* for time n, $n = 0, 1, 2, \ldots$, under the hypothesis $\mathfrak{S}_n$ if control is performed with the strategy ϑ by

$$V_{n,\vartheta} := \boldsymbol{E}_\vartheta(R_{n+1,\vartheta}|\mathfrak{S}_n) \tag{11}$$

with the realisations

$$v_{n,\vartheta}(\mathfrak{s}_n) := \boldsymbol{E}_\vartheta(R_{n+1,\vartheta}|\mathfrak{S}_n = \mathfrak{s}_n). \tag{11'}$$

In case $N < \infty$ we put

$$v_{N,\vartheta}(\mathfrak{s}_N) := 0 \qquad \text{for all} \quad \mathfrak{s}_N \in \mathscr{S}^{N+1}, \ \vartheta \in \Theta.$$

For $\vartheta \sim h_n$, $\mathfrak{s}_n = \mathfrak{s}_n(h_n)$ we still denote

$$v_{n,\vartheta}(\mathfrak{s}_n) =: w_{n,\vartheta}(h_n). \tag{12}$$

Now our aim function (4) may be represented as *mean total cost* for the strategy ϑ and for each $n \geqslant 0$ decomposed in the form (compare (7), (3), (9), (11))

$$v_\vartheta = \boldsymbol{E}_\vartheta(K_{n,\vartheta}) + \boldsymbol{E}_\vartheta(V_{n,\vartheta}); \tag{13}$$

because of (11') we obtain for $n = 0$ especially

$$v_\vartheta = \sum_s p_0(s)\, \boldsymbol{E}_\vartheta(R_{1,\vartheta}|X_0 = s) = \sum_s p_0(s)\, v_{0,\vartheta}(s). \tag{14}$$

For an arbitrary strategy $\vartheta = (\delta_k) \in \Theta$ and $0 \leqslant m < n < N$ with $m, n \in \boldsymbol{N}$ we define the (*strategy-*) *section*

$$\vartheta_{[m,n]} := (\delta_m, \delta_{m+1}, \ldots, \delta_n). \tag{15}$$

$$\Theta_n(\vartheta) := \{\eta \in \Theta\colon\ \eta_{[0,n]} = \vartheta_{[0,n]}\} \tag{16}$$

signifies the set of all strategies of Θ, which coincide with ϑ till step n; the elements of $\Theta_n(\vartheta)$ are called $(n+1)$-*continuations* of ϑ. In addition we put $\Theta_{-1}(\vartheta) := \Theta$ for arbitrary $\vartheta \in \Theta$.

Departing from

$$V_{n,\eta} = \boldsymbol{E}_\eta(R_{n+1,\eta}|\mathfrak{S}_n),$$

compare (11), we consider for $\vartheta \in \Theta$ and $n \geqslant 0$ the random variables

$$V^*_{\vartheta|n} := \inf_{\eta \in \Theta_{n-1}(\vartheta)} V_{n,\vartheta},$$

especially for $n = 0$ we put because of $\Theta_{-1}(\vartheta) = \Theta$ (17)

$$V_{\vartheta|0}^{*} = \inf_{\eta \in \Theta} V_{0,\eta} =: V_0^{*}.$$

As to the existence of these random variables we take into account, that the state-space is supposed to be countable. $V_{\vartheta|0}^{*}$ is obviously independent of ϑ. For the notation of the corresponding realizations we write for $\omega \in \Omega$ with $a_n \omega = \mathfrak{s}_n$, $n \geqslant 0$ instead of $V_{\vartheta|n}^{*}(\omega)$ likewise $v_{\vartheta|n}^{*}(\mathfrak{s}_n)$, and it follows by definition

$$v_{\vartheta|n}^{*}(\mathfrak{s}_n) = \inf_{\eta \in \Theta_{n-1}(\vartheta)} v_{n,\vartheta}(\mathfrak{s}_n) \qquad \text{for} \quad \boldsymbol{P}_\vartheta(\mathfrak{S}_n = \mathfrak{s}_n) > 0$$

and especially for $n = 0$ (17′)

$$v_{\vartheta|0}^{*}(s_0) = v_0^{*}(s_0) = \inf_{\eta \in \Theta} v_{0,\eta}(s_0) \qquad \text{for} \quad p_0(s_0) > 0.$$

In case of $\boldsymbol{P}_\vartheta(\mathfrak{S}_n = \mathfrak{s}_n) = 0$ we put

$$v_{\vartheta|n}^{*}(\mathfrak{s}_n) := 0,$$

and in consequence we define

$$v_0^{*}(s_0) := 0 \qquad \text{for} \quad p_0(s_0) = 0.$$

The existence of the precedingly introduced random variables is ensured by the following assumption

(V): For all $\vartheta \in \Theta$ and $n \geqslant 0$, $n \in \boldsymbol{N}$ there exist $\boldsymbol{P}_\vartheta$-a.s. $V_{n,\vartheta}$ with $\boldsymbol{E}_\vartheta V_{n,\vartheta}^{+} < \infty$. (18)

As a consequence of (V) we have

$$\boldsymbol{E}_\vartheta V_{n,\vartheta} < +\infty.$$

6

DEFINITION. (a) For $\vartheta \in \Theta$, $n \geqslant 0$ and $\mathfrak{s}_n \in \mathscr{S}^{n+1}$ with $\boldsymbol{P}_\vartheta(\mathfrak{S}_n = \mathfrak{s}_n) > 0$ we denote

$$\mathscr{S}_{n+1}^{+}(\vartheta, \mathfrak{s}_n) := \{s \in \mathscr{S} \colon \; p_{n+1,\vartheta}(s|\mathfrak{s}_n) > 0\}.$$

If exists always, this means for each choice of the precedingly quoted symbols and in addition for every positive ε, and for each two strategies $\vartheta_1, \vartheta_2 \in \Theta_n(\vartheta)$ another strategy $\vartheta \in \Theta_n(\vartheta)$ with

$$v_{n+1,\vartheta'}(\mathfrak{s}_n, s) \leqslant \min\left(v_{n+1,\vartheta_1}(\mathfrak{s}_n, s),\, v_{n+1,\vartheta_2}(\mathfrak{s}_n, s)\right) + \varepsilon$$

for all $s \in \mathscr{S}_{n+1}^{+}(\vartheta, \mathfrak{s}_n)$, then we call Θ *weakly complete* (*relative to minorization by pairs*). We symbolize this property by (V_p'). If there exists for each two strategies $\vartheta_1, \vartheta_2 \in \Theta_n(\vartheta)$ always a strategy $\vartheta' \in \Theta_n(\vartheta)$ with

$$v_{n+1,\vartheta'}(\mathfrak{s}_n, s) \leqslant \min(v_{n+1,\vartheta_1}(\mathfrak{s}_n, s), v_{n+1,\vartheta_2}(\mathfrak{s}_n, s))$$

for all $s \in \mathscr{S}^+_{n+1}(\vartheta, \mathfrak{s}_n)$, then Θ is called *complete* (*relative to minorization by pairs*). In symbols: (V_p).

(b) The set Θ of admissible strategies is called *strongly complete* (*relative to uniformly optimal continuation*), if there exists to each $n \geqslant 0$ and to each $\vartheta \in \Theta$ a strategy $\vartheta^{*n} \in \Theta_{n-1}(\vartheta)$, so that

$$v_{n,\vartheta^{*n}}(\mathfrak{s}_n) = v^*_{\vartheta|n}(\mathfrak{s}_n)$$

is valid for each state-n-history $\mathfrak{s}_n$ being essential for ϑ; in symbols (V_{un}). In case $n \geqslant 1$ we call ϑ^{*n} a (*uniformly*) *optimal n-continuation of* ϑ. In case $n = 0$ we have accordingly to (17′)

$$v_{0,\vartheta^{*0}}(s_0) = v^*_0(s_0)$$

for all s_0 with $p_0(s_0) > 0$, and in consequence of (14) and (5) ϑ^{*0} is an optimal strategy. By means of the precedingly introduced symbols (V_{un}) implies (V_p) and (V_p) implies (V_p').

7

LEMMA. *Let be condition* (V) – see (18) – *fulfilled and we suppose the set* Θ *of admissible strategies to be weakly complete* (compare Definition 6(a)), *then to each* $\varepsilon > 0$, *to each* $\vartheta \in \Theta$ *and to each state-n-history* $\mathfrak{s}_n$, $n \geqslant 0$, *being essential for* ϑ, *there exists a strategy* $\eta \in \Theta_n(\vartheta)$, *such that*

$$\boldsymbol{E}_\vartheta(V_{n+1,\eta}|\mathfrak{S}_n = \mathfrak{s}_n) \leqslant \boldsymbol{E}_\vartheta(V^*_{\vartheta|n+1}|\mathfrak{S}_n = \mathfrak{s}_n) + \varepsilon. \tag{19}$$

Proof. To each state-n-history $\mathfrak{s}_n$ being essential for the strategy $\vartheta \in \Theta$, $\mathfrak{s}_n \in \{\mathscr{S}^{n+1}\}^+(\vartheta)$, $n \geqslant 0$, there exists a countable set $T_{\vartheta,n}(\mathfrak{s}_n) \in \Theta_{n-1}(\vartheta)$ with

$$\inf_{\eta \in T_{n,\vartheta}(\mathfrak{s}_n)} v_{n,\eta}(\mathfrak{s}_n) = v^*_{\vartheta|n}(\mathfrak{s}_n).$$

As $\mathscr{S}$ is countable, it follows that $\mathscr{S}^{n+1}$ and consequently the subset $\{\mathscr{S}^{n+1}\}^+(\vartheta)$ of $\mathscr{S}^{n+1}$ and further

$$\bigcup_{\mathfrak{s}_n = \{\mathscr{S}^{n+1}\}^+(\vartheta)} T_{\vartheta,n}(\mathfrak{s}_n) := T_{\vartheta,n}$$

are also countable, and we have the equation $V^*_{\vartheta|n} = \inf_{\eta \in T_{\vartheta,n}} V_{n,\vartheta}$, $\boldsymbol{P}_\vartheta$-a.e.

Now let be $T_{\vartheta,n+1} =: (\hat{\eta}_k)$, $k = 1, 2, \ldots$, a countable subset of $\Theta_n(\vartheta)$ existing by the preceding considerations with

$$V^*_{\vartheta|n+1} = \inf_{\hat{\eta} \in T_{\vartheta,n+1}} V_{n+1,\hat{\eta}} \qquad \boldsymbol{P}_\vartheta\text{-a.e.}$$

We still put $\mathscr{S}' := \{s \in \mathscr{S}: p_{n+1,\vartheta}(s|\mathfrak{s}_n) > 0\} = \mathscr{S}^+_{n+1}(\vartheta, \mathfrak{s}_n)$ and we write in abbreviation

$$V_{n+1,\eta} =: V_\eta, \qquad V^*_{\vartheta|n+1} =: V^*_\vartheta,$$

$$v_{n+1,\eta}(\mathfrak{s}_n, s) =: v_\eta(\mathfrak{s}_n, s) \quad \text{and} \quad v^*_{\vartheta|n+1}(\mathfrak{s}_n, s) =: v^*_\vartheta(\mathfrak{s}_n, s).$$

To given $\varepsilon > 0$ we define departing from $\eta_1 := \hat{\eta}_1$ and by virtue of the weak completeness of Θ a sequence (η_k), $k = 1, 2, \ldots$ with $\eta_k \in \Theta_n(\vartheta)$ and

$$v_{\eta_j}(\mathfrak{s}_n, s) \leqslant \min\left(v_{\hat{\eta}_j}(\mathfrak{s}_n, s), v_{\eta_{j-1}}(\mathfrak{s}_n, s)\right) + \varepsilon_j, \qquad s \in \mathscr{S}', \ j \geqslant 2,$$

whereby $\varepsilon_j = \varepsilon/2^j$. With

$$\min\left(v_{\hat{\eta}_1}(\mathfrak{s}_n, s), \ldots, v_{\hat{\eta}_j}(\mathfrak{s}_n, s)\right) =: \varphi_j(\mathfrak{s}_n, s)$$

it follows

$$v_{\eta_j}(\mathfrak{s}_n, s) \leqslant \varphi_j(\mathfrak{s}_n, s) + \varepsilon_2 + \ldots + \varepsilon_j = \varphi_j(\mathfrak{s}_n, s) + \tfrac{1}{2}\varepsilon, \qquad s \in \mathscr{S}', \tag{$\circ$}$$

As by definition of $(\hat{\eta}_k)$, $k = 1, 2, \ldots$, the sequence $\left(\varphi_j(\mathfrak{s}_n, s)\right)$, $j = 1, 2, \ldots$ for each $s \in \mathscr{S}'$ is monotonously decreasing convergent to $v^*_\vartheta(\mathfrak{s}_n, s)$, we have by a well-known integral-theorem an integer j_0, such that with $\mathfrak{S}_{n+1} = (X_0, \ldots, X_{n+1})$ and $\varphi_{j_0}(\mathfrak{S}_{n+1}) = \Phi_{j_0}$

$$\boldsymbol{E}_\vartheta(\Phi_{j_0}|\mathfrak{S}_n = \mathfrak{s}_n) \leqslant \boldsymbol{E}_\vartheta(V^*_\vartheta|\mathfrak{S}_n = \mathfrak{s}_n) + \tfrac{1}{2}\varepsilon. \tag{$\circ\circ$}$$

By ($\circ$) and ($\circ\circ$) now follows

$$\boldsymbol{E}_\vartheta(V_{j_0}|\mathfrak{S}_n = \mathfrak{s}_n) \leqslant \boldsymbol{E}_\vartheta(V^*_\vartheta|\mathfrak{S}_n = \mathfrak{s}_n) + \varepsilon$$

after which with $\eta_{j_0} =: \eta$ we obtain the assertion. ∎

For the proof of the optimality-equation given later on, we add a transformation of (19). For $\eta \in \Theta_n(\vartheta)$ we have

$$p_{n+1,\vartheta}(s|\mathfrak{s}_n) = p_{n+1,\eta}(s|\mathfrak{s}_n),$$

and from

$$\boldsymbol{P}_\vartheta(X_{n+1} = s|\mathfrak{S}_n = \mathfrak{s}_n) = p_{n+1,\vartheta}(s|\mathfrak{s}_n),$$

compare (2), it follows

$$\boldsymbol{E}_\vartheta(V_{n+1,\eta}|\mathfrak{S}_n = \mathfrak{s}_n) = \sum_s v_{n+1,\eta}(\mathfrak{s}_n, s)\, p_{n+1,\eta}(s|\mathfrak{s}_n)$$

and correspondingly

$$\boldsymbol{E}_\vartheta(V^*_{\vartheta|n+1}|\mathfrak{S}_n = \mathfrak{s}_n) = \sum_s v^*_{\vartheta|n+1}(\mathfrak{s}_n, s)\, p_{n+1,\eta}(s|\mathfrak{s}_n).$$

Therefore (19) is equivalent to

$$\sum_s v_{n+1,\eta}(\mathfrak{s}_n, s)\, p_{n+1,\eta}(s|\mathfrak{s}_n) \leqslant \sum_s v^*_{\vartheta|n+1}(\mathfrak{s}_n, s)\, p_{n+1,\eta}(s|\mathfrak{s}_n) + \varepsilon. \tag{19'}$$

8

DEFINITION. To a given set Θ of admissible strategies we introduce for $n \geqslant 0$ the following sets:

$$E_n := \{\zeta | \mathscr{H}_n \to \mathscr{D} : \exists \vartheta = (\delta_k) \in \Theta \text{ with } \zeta = \delta_n\}, \tag{20}$$

the set of all admissible decision functions at time n;

$$\mathscr{D}_n(h) := \{d \in \mathscr{D} : \exists \zeta \in E_n \text{ with } \zeta(h) = d\}, \qquad h \in \mathscr{H}_n, \tag{21}$$

the set of all admissible decisions at time n to the given n-history h;

$$\Delta_n := \{\zeta | \mathscr{H}_n \to \mathscr{D} : \zeta(h) \in \mathscr{D}_n(h),\ h \in \mathscr{H}_n\}, \tag{22}$$

the set of all functions on $\mathscr{H}_n$ with values in the corresponding sets of decisions.

Evidently we have $E_n \subset \Delta_n$, and it follows

$$\Theta \subset \mathop{\times}_{0 \leqslant n < N} E_n \subset \mathop{\times}_{0 \leqslant n < N} \Delta_n. \tag{23}$$

We call Θ (with respect to the corresponding set of decision-functions) *combinable*, if

$$\Theta = \mathop{\times}_{0 \leqslant n < N} E_n \tag{24}$$

complete, if

$$E_n = \Delta_n, \qquad n \geqslant 0, \tag{25}$$

and *maximal*, if Θ is both combinable and complete, i.e.,

$$\Theta = \mathop{\times}_{0 \leqslant n < N} \Delta_n. \tag{26}$$

9

Remark. If Θ is combinable, the admissible strategies may be formed by an arbitrary combination of decision functions $\delta_n \in E_n$ to the individual points of time $n = 0, \ldots$ If $\vartheta^1 = (\delta_k^1)$ and $\vartheta^2 = (\delta_k^2)$ are two strategies of an arbitrary set Θ of admissible strategies, we denote by $\vartheta^1_{[n]}\vartheta^2$ the strategy $\vartheta = (\delta_k)$ formed by combination of ϑ^1 and ϑ^2, for which $\delta_k = \delta_k^1$ if $k \leqslant n$ and $\delta_k = \delta_k^2$ if $k > n$. According to (24) the set Θ is combinable, iff to each two strategies ϑ^1 and ϑ^2 from Θ and for arbitrary $n \geqslant 0$ also the strategy $\vartheta^1_{[n]}\vartheta^2$ is contained in Θ.

In case of a complete set of admissible strategies at each point of time $n \geqslant 0$ the totality of functions on $\mathscr{H}_n$ with values in $\mathscr{D}_n(h)$, $h \in \mathscr{H}_n$, is used as set of admissible decision functions, but in difference to the case of a combinable Θ not each possible combination (δ_n) of decision functions δ_n, $n \geqslant 0$, is necessarily an admissible strategy.

A rather simple but for many real situations fitting case is given by maximal sets of admissible strategies. We demonstrate by the following examples that the inclusions figuring in (23) may in general not be replaced by equalities.

10

EXAMPLES. (a) Let be $N = 2$, $\mathscr{S}$ respectively $\mathscr{D}$ denote the sets of states and decisions and Δ_0 and Δ_1 the sets of all functions from $\mathscr{H}_0$ respectively $\mathscr{H}_1$ in $\mathscr{D}$. For any special $\delta \in \Delta_1$ let be $\Delta_1' := \Delta_1 \setminus \{\delta\}$. Then the set of strategies $\Theta := \Delta_1 \times \Delta_1'$ is combinable, but not complete. Not combinable sets of admissible strategies are arising for instance, if single strategies are removed. In connection with the theory of optimal stopping we still consider the following example. In the dynamical system 1 let be $N = \infty$ and $\mathscr{D} = \{0, 1\}$. We define a strategy $\vartheta = (\delta_k)$ to be admissible, i.e. $\vartheta \in \Theta$, if for each $\omega \in \Omega$ there exists an integer $n < \infty$ with $\delta_n(a_n \omega) = 1$. For all $k \in N_0$ and $h_k \in \mathscr{H}_k$ we have $\mathscr{D}_k(h_k) = \{0, 1\}$. Therefore each strategy $\vartheta = (\delta_k)$, for which for arbitrary $\omega \in \Omega$ and every $n \in N_0$ we always have $\delta_n(h_n(a_n \omega)) = 0$, is contained in $\bigtimes_{0 \leqslant n < N} \Delta_n$, but by supposition not in Θ. Hence Θ is not combinable.

(b) Let be $\Theta = (\Delta_0' \times \Delta_1') \cup (\Delta_0'' \times \Delta_1'')$ and $\Delta_i = \Delta_i' \cup \Delta_i''$, $\Delta_i' \cap \Delta_i'' = \emptyset$ with $\Delta_i' \neq \emptyset$, $\Delta_i'' \neq \emptyset$, $i = 1, 2$. Evidently Θ is complete, but because of $(\Delta_0' \times \Delta_1'') \cap \Theta = (\Delta_0'' \times \Delta_1') \cap \Theta = \emptyset$ not combinable.

In literature sets of admissible strategies are frequently introduced in this way, compare for instance [2], that at first are given for each point of time n, $n \geqslant 0$, in dependance of each n-history h the set of all admissible decisions. We denote these sets again by $\mathscr{D}_n(h)$, however regarding that for the moment strategies and decision functions are not yet present. Then for $n \geqslant 0$ the totality of all functions δ_n with $\delta_n(h) \in \mathscr{D}_n(h)$ for $h \in \mathscr{H}_n$ is introduced as set Δ_n of all admissible decision functions at time n. The set of all admissible strategies is then defined by $\Theta = \bigtimes_{0 \leqslant n < N} \Delta_n$, i.e., $\Theta = \{\vartheta = (\delta_n): \delta_n \in \Delta_n, n \geqslant 0\}$, which is accordingly to (26) maximal by definition.

Let F be an arbitrary function from $\mathscr{D}$ in $\boldsymbol{R}$. For $n \geqslant 0$, $h \in \mathscr{H}_n$, $\vartheta \in \Theta$ with $\vartheta \sim h$ we consider

$$\inf_{\eta = (\zeta_k) \in \Theta_{n-1}(\vartheta)} F(\zeta_n(h)) =: F^*. \tag{27}$$

If Θ is combinable it follows according to (24)

$$F^* = \inf_{\zeta_n \in E_n} F(\zeta_n(h)), \tag{28}$$

and if Θ is maximal, it follows according to (26)

$$F^* = \inf_{d \in \mathscr{D}_n(h)} F(d). \tag{29}$$

The preceding equations signify, that the infimum, formulated in (27), which relates to all strategies, which are n-continuations of ϑ, can be reduced in case of a combinable Θ to the decision functions of E_n, and in case of a maximal Θ to the decisions of the set $\mathscr{D}_n(h)$. This will be used for appropriate formulations of the optimality equation.

11

LEMMA. *Let be h_n a n-history and ϑ a strategy in Θ with $\vartheta \sim h_n$. If Θ is combinable, then for*

$$w_{n,\eta}(h_n) = E_\eta(R_{n+1,\eta} | \mathfrak{S}_n = \mathfrak{s}_n)$$

with $h_n = h_n^\eta(\mathfrak{s}_n)$ – compare (11′) and (11″) – it follows

$$\inf_{\eta \in \Theta_{n-1}(\vartheta)} w_{n,\eta}(h_n) = \inf_{\eta \sim h_n} w_{n,\eta}(h_n). \tag{30}$$

Proof. We put

$$\inf_{\eta \in \Theta_{n-1}(\vartheta)} w_{n,\eta}(h_n) =: A, \qquad \inf_{\eta \sim h_n} w_{n,\eta}(h_n) =: B.$$

(a) In consequence of $\eta \in \Theta_{n-1}(\vartheta) \Rightarrow \eta \sim h_n$ we have immediately $A \geqslant B$.

(b) To each $\varepsilon > 0$ there exists a $\eta' \sim h_n$ with

$$w_{n,\eta'}(h_n) < B + \varepsilon.$$

As Θ is by supposition combinable, the strategy $\vartheta_{(n-1]}\eta' =: \vartheta'$ also belongs to Θ, and we have $\vartheta' \in \Theta_{n-1}(\vartheta)$ and hence $\vartheta' \sim h_n$. Because of $h_n := h_n^\vartheta(\mathfrak{s}_n) = h_n^{\eta'}(\mathfrak{s}_n) = h_n^{\vartheta'}(s_n)$ then follows the equation $R_{n+1,\vartheta'} = R_{n+1,\eta'}$ almost sure with respect to the measures $P_{\vartheta'}$ and $P_{\eta'}$. Further we have

$$w_{n,\vartheta'}(h_n) = E_{\vartheta'}(R_{n+1,\vartheta'} | \mathfrak{S}_n = \mathfrak{s}_n) = E_{\eta'}(R_{n+1,\eta'} | \mathfrak{S}_n = \mathfrak{s}_n) = w_{n,\eta'}(h_n),$$

as the expectations are independent of the process till n, and the strategies coincide after that. Hence follows

$$w_{n,\vartheta'}(h_n) < B + \varepsilon$$

for arbitrary $\varepsilon > 0$ and from this $A \leqslant B$.

From (a) and (b) the statement of the lemma follows. ∎

For $h_n \in \mathscr{H}_n$ we denote

$$\inf_{\eta \sim h_n} w_{n,\eta}(h_n) =: w_n^*(h_n). \tag{31}$$

12

LEMMA. *If Θ is maximal* (see Definition 6), *then Θ is complete (relative to minorization by pairs).*

Proof. Let be $\vartheta = (\delta_k) \in \Theta$, $n \geqslant 0$, $\mathfrak{s}_n \in \{\mathscr{S}^{n+1}\}^+(\vartheta)$ and $\vartheta_1 = (\delta_k^1) \in \Theta_n(\vartheta)$ and $\vartheta_2 = (\delta_k^2) \in \Theta$ arbitrarily given. As Θ is maximal by supposition, Θ contains the strategy $\vartheta' = (\delta_k')$, defined by $\delta_k' = \delta_k$ if $k \leqslant n$, and for $k \geqslant n+1$

$$\delta_k'(h^{\vartheta'}(\mathfrak{s}_k)) = \begin{cases} \delta_k^1(h^{\vartheta_1}(\mathfrak{s}_k)) & \text{if} \quad v_{n+1,\vartheta_1}(\mathfrak{s}_n, s) \leqslant v_{n+1,\vartheta_2}(\mathfrak{s}_n, s), \\ \delta_k^2(h^{\vartheta_2}(\mathfrak{s}_k)) & \text{if} \quad v_{n+1,\vartheta_1}(\mathfrak{s}_n, s) > v_{n+1,\vartheta_2}(\mathfrak{s}_n, s), \end{cases}$$

at which $a_{n+1}\,\mathfrak{s}_k = (\mathfrak{s}_n, s)$. From this especially for $s \in \mathscr{S}$ with $p_{n+1,\vartheta}(s|\mathfrak{s}_n) > 0$ we obtain the relation

$$v_{n+1,\vartheta'}(\mathfrak{s}_n, s) \leqslant \min\big(v_{n+1,\vartheta_1}(\mathfrak{s}_n, s),\, v_{n+1,\vartheta_2}(\mathfrak{s}_n, s)\big),$$

whence the statement of the lemma follows. ∎

13. A recurrence formula

In preparation of the optimality equation we next formulate a simple recurrence relation between mean rest cost to neighbouring points of time. Taking notice of (6), (7), (8), (10) and (11) we obtain from

$$R_{n+1,\vartheta} = C_{n+1,\vartheta} + R_{n+2,\vartheta}$$

and

$$\begin{aligned} \boldsymbol{E}_\vartheta(R_{n+2,\vartheta}|\mathfrak{S}_n = \mathfrak{s}_n) &= \boldsymbol{E}_\vartheta\big(\boldsymbol{E}_\vartheta(R_{n+2,\vartheta}|\mathfrak{S}_{n+1})|\mathfrak{S}_n = \mathfrak{s}_n\big) \\ &= \boldsymbol{E}_\vartheta(V_{n+1,\vartheta}|\mathfrak{S}_n = \mathfrak{s}_n) \qquad \text{for} \quad n = 0, 1, \dots \end{aligned}$$

the equation

$$\begin{aligned} v_{n,\vartheta}(\mathfrak{s}_n) &= \boldsymbol{E}_\vartheta(R_{n+1,\vartheta}|\mathfrak{S}_n = \mathfrak{s}_n) \\ &= \boldsymbol{E}_\vartheta(C_{n+1,\vartheta} + R_{n+2,\vartheta}|\mathfrak{s}_n = \mathfrak{s}_n) \\ &= \boldsymbol{E}_\vartheta(C_{n+1,\vartheta}|\mathfrak{S}_n = \mathfrak{s}_n) + \boldsymbol{E}_\vartheta(R_{n+2,\vartheta}|\mathfrak{S}_n = \mathfrak{s}_n) \\ &= \boldsymbol{E}_\vartheta(C_{n+1,\vartheta}|\mathfrak{S}_n = \mathfrak{s}_n) + \boldsymbol{E}_\vartheta(V_{n+1,\vartheta}|\mathfrak{S}_n = \mathfrak{s}_n). \end{aligned} \tag{32}$$

Now we have

$$\boldsymbol{E}_\vartheta(V_{n+1,\vartheta}|\mathfrak{S}_n = \mathfrak{s}_n) = \sum_s p_{n+1,\vartheta}(s|\mathfrak{s}_n)\, v_{n+1,\vartheta}(\mathfrak{s}_n, s).$$

Hence from (32) and (10) we obtain

$$v_{n,\vartheta}(\mathfrak{s}_n) = \sum_s p_{n+1,\vartheta}(s|\mathfrak{s}_n)\,[c_{n+1,\vartheta}(\mathfrak{s}_n, s) + v_{n+1,\vartheta}(\mathfrak{s}_n, s)], \quad n = 0, 1, 2, \dots \tag{33}$$

For $\vartheta = (\delta_k) \sim h_n$, $\mathfrak{s}_n = \mathfrak{s}_n(h_n)$ it follows because of (1), (10), (10') and (12) from (33)

$$w_{n,\vartheta}(h_n) = \mathfrak{S}_n(h_n, d_n) + \sum_s p_{n+1}(s|h_n, d_n) w_{n+1,\vartheta}(h_n, d_n, s),$$
$$h_n \in \mathscr{H}_n,\ \delta_n(h_n) = d_n,\ n = 0, 1, 2, \ldots \qquad (33')$$

The formula expresses the simple but basic fact, that mean rest cost for a special point of time yield recurrently by addition of the next mean step cost and the mean rest cost for the following point of time.

14. Optimality-equation

THEOREM. *Let ϑ be an admissible strategy and $\mathfrak{s}_n$ be a state-n-history being essential for ϑ, $n \geqslant 0$. We assume that in addition to the condition* (V) (see (18)) *Θ shall be weakly complete relative to minorization by pairs* (see Definition 6). *Then the following optimality-equation holds* (for notation we refer to (1), (8), and (17')

$$v^*_{\vartheta|n}(\mathfrak{s}_n) = \inf_{\eta \in \Theta_{n-1}(\vartheta)} \left\{\sum_s p_{n+1,\eta}(s|\mathfrak{s}_n)\left[c_{n+1,\eta}(\mathfrak{s}_n, s) + v^*_{\eta|n+1}(\mathfrak{s}_n, s)\right]\right\}. \qquad (34)$$

(We take notice on our definition $\Theta_{-1}(\vartheta) := \Theta$ made after formula (16).)

Proof. (a) For an arbitrary strategy $\vartheta' \in \Theta_{n-1}(\vartheta)$, $n \geqslant 0$, we have in consequence of (33)

$$\begin{aligned} v_{n,\vartheta'}(\mathfrak{s}_n) &= \sum_s p_{n+1,\vartheta'}(s|\mathfrak{s}_n)\left[c_{n+1,\vartheta'}(\mathfrak{s}_n, s) + v_{n+1,\vartheta'}(\mathfrak{s}_n, s)\right] \\ &\geqslant \sum_s p_{n+1,\vartheta'}(s|\mathfrak{s}_n)\left[c_{n+1,\vartheta'}(\mathfrak{s}_n, s) + v^*_{\vartheta'|n+1}(\mathfrak{s}_n, s)\right] \\ &\geqslant \inf_{\eta \in \Theta_{n-1}(\vartheta)} \left\{\sum_s p_{n+1,\eta}(s|\mathfrak{s}_n)\left[c_{n+1,\eta}(\mathfrak{s}_n, s) + v^*_{\eta|n+1}(\mathfrak{s}_n, s)\right]\right\}. \qquad (\circ) \end{aligned}$$

(b) If $v^*_{\vartheta|n}(\mathfrak{s}_n) = -\infty$, then to every positive M exists a strategy $\vartheta' \in \Theta_{n-1}(\vartheta)$ with $v_{n,\vartheta'}(\mathfrak{s}_n) < -M$, hence by ($\circ$) follows

$$-M > \inf_{\eta \in \Theta_{n-1}(\vartheta)} \left\{\sum_s p_{n+1,\eta}(s|\mathfrak{s}_n)\left[c_{n+1,\eta}(\mathfrak{s}_n, s) + v^*_{\eta|n+1}(\mathfrak{s}_n, s)\right]\right\}.$$

From this we immediately obtain (34).

(c_1) Let be $v^*_{\vartheta|n}(\mathfrak{s}_n) > -\infty$. Then to arbitrary positive ε we have a strategy $\vartheta' \in \Theta_{n-1}(\vartheta)$ with

$$v_{n,\vartheta'}(\mathfrak{s}_n) < v^*_{\vartheta|n}(\mathfrak{s}_n) + \varepsilon,$$

so that from ($\circ$)

$$v^*_{\vartheta|n}(\mathfrak{s}_n) + \varepsilon > \inf_{\eta \in \Theta_{n-1}(\vartheta)} \left\{\sum_s p_{n+1,\eta}(s|\mathfrak{s}_n)\left[c_{n+1,\eta}(\mathfrak{s}_n, s) + v^*_{\eta|n+1}(\mathfrak{s}_n, s)\right]\right\}$$

follows.

(c_2) According to (19′) to each $\eta \in \Theta_{n-1}(\vartheta)$ and to arbitrary positive ε exists a strategy $\eta' \in \Theta_n(\eta)$, such that

$$\sum_s p_{n+1,\eta'}(s|\mathfrak{s}_n)\, v_{n+1,\eta'}(\mathfrak{s}_n, s) \leqslant \sum_s p_{n+1,\eta'}(s|\mathfrak{s}_n)\, v^*_{\eta|n+1}(\mathfrak{s}_n, s) + \varepsilon.$$

Hence from (33) and according to the equations

$$p_{n+1,\eta'}(s|\mathfrak{s}_n) = p_{n+1,\eta}(s|\mathfrak{s}_n) \text{ and } c_{n+1,\eta'}(\mathfrak{s}_n, s) = c_{n+1,\eta}(\mathfrak{s}_n, s)$$

it follows

$$v^*_{\vartheta|n}(\mathfrak{s}_n) \leqslant v_{n,\eta'}(\mathfrak{s}_n) = \sum_s p_{n+1,\eta'}\,[c_{n+1,\eta'}(\mathfrak{s}_n, s) + v_{n+1,\eta'}(\mathfrak{s}_n, s)]$$
$$\leqslant \sum_s p_{n+1,\eta}(s|\mathfrak{s}_n)\,[c_{n+1,\eta}(\mathfrak{s}_n, s) + v^*_{\eta|n+1}(\mathfrak{s}_n, s)] + \varepsilon$$

for each $\eta \in \Theta_{n-1}(\vartheta)$, and consequently

$$v^*_{\vartheta|n}(\mathfrak{s}_n) \leqslant \inf_{\eta \in \Theta_{n-1}(\vartheta)} \Big\{\sum_s p_{n+1,\eta}(s|\mathfrak{s}_n)\,[c_{n+1,\eta}(\mathfrak{s}_n, s) + v^*_{\eta|n+1}(\mathfrak{s}_n, s)]\Big\} + \varepsilon.$$

As ε in (c_1) and (c_2) is an arbitrary positive number, the validity of the optimality equation (34) follows. ■

We deduce the special form of the optimality equation for the case, that the set Θ of the admissible strategies is combinable respectively maximal; see Definition 8. Firstly we have for $\eta = (\beta_k)$ according to (1), (8) with $h_n = h^\eta_n(\mathfrak{s}_n)$, $\mathfrak{s}_n \in \mathscr{S}^{n+1}$, $n \geqslant 0$, $s \in \mathscr{S}$

$$p_{n+1,\eta}(s|\mathfrak{s}_n) = p_{n+1}(s|h_n, \beta_n(h_n)), \qquad c_{n+1,\eta}(\mathfrak{s}_n, s) = c_{n+1}(h_n, \beta_n(h_n), s)$$

and else for combinable Θ, compare (12), (30) and (31),

$$v^*_{\eta|n+1}(\mathfrak{s}_n, s) = w^*_{n+1}(h_n, \beta_n(h_n), s).$$

Hence from (34) according to (24) we obtain for $n = 0, 1, \ldots$ and each essential n-history h_n in case of a combinable set Θ of admissible strategies the optimality equation

$$w^*_n(h_n) = \inf_{\beta_n \in E_n} \Big\{\sum_s p_{n+1}(s|h_n, \beta_n(h_n))\,[c_{n+1}(h_n, \beta_n(h_n), s) + w^*_{n+1}(h_n, \beta_n(h_n), s)]\Big\}.$$

If Θ is moreover maximal, it follows from (35) according to (29)

$$w^*_n(h_n) = \inf_{d \in \mathscr{D}_n(h_n)} \Big\{\sum_s p_{n+1}(s|h_n, d)\,[c_{n+1}(h_n, d, s) + w^*_{n+1}(h_n, d, s)]\Big\}. \quad (36)$$

15

Remark. The idea for establishing of Lemma 7 was suggested by [1].

References

[1] T. Bojdecki, *Martyngały z czasem dyskretnym. Zarys teorii i przykłady zastosowań*, Wydawnictwa Uniwersytetu Warszawskiego, Warszawa 1977.

[2] K. Hinderer, *Foundations of non-stationary dynamic programming with discrete time-parameter*, Lecture Notes in Oper. Res. and Math. Syst. **33**, Springer-Verlag, Berlin 1970.

Presented to the semester
Sequential Methods in Statistics
September 7–December 11, 1981